6

Representing Cognition in Games and Simulations

J. D. FLETCHER AND JOHN E. MORRISON

Institute for Defense Analyses

Developers of games and simulations are striving to increase realism in both the appearance and the behavior of their computer-generated characters. The physical appearance of characters in games and simulations enhances participants' sense of presence and immersion, but the way the characters behave may matter more (Bailenson et al., 2005; Garau, Slater, Bee, & Sasse, 2001). Interest in imbuing computer-generated characters with human-like behavior is present and growing.

This interest approaches an imperative when it comes to generating friendly, opposition, and neutral characters for simulations used in military training. One side will implement a tactic that the other side will successfully counter, requiring the first side to adjust their tactics, which will affect what the other side does next, and so on. This process creates immersing, problem-solving, decision-making environments that can test the limits of both human and machine cognition. It mimics the poorly structured, rapidly shifting, ill-defined, and time-constrained environments that are typical of real-world problem solving and decision making (Cannon-Bowers, Salas, & Pruitt, 1996). Interest in cognitively realistic characters may be equally imperative in simulating environments for nonmilitary applications, such as those developed for problem solving and decision making in urban planning, economic management, and of course, multiuser fantasy games.

This chapter reviews models that are available for representing human cognitive behavior in games and simulations. It identifies and briefly discusses those models that seem particularly appropriate for representing and assessing human problem-solving capabilities.

Games and Simulations

Computer-based simulations appeared with the first computers that could support them and, for that matter, were a prime motivator for their development (Goldstine, 1972). Simulations intended for education and training appeared almost as soon. Perhaps the first example of computer-based simulation training grew out of the Air Force SAGE (Semi-Automatic Ground Environment) system (Rowell & Streich, 1964). The early 1950s Whirlwind I project demonstrated that computers using radar data to track aircraft could serve as an early warning air defense system for the entire North American continent. Based on this evidence and motivated by the perceived exigencies of the Cold War, the Air Force quickly responded by building SAGE, a system of 20 computer-based, linked direction centers for tracking aircraft and controlling aircraft interceptions.

On the human factors side, SAGE was a large, geographically dispersed system requiring intense human-computer interactions, complex and frequent decision making, and close cooperation among many operators working in relative psychological isolation (Rowell & Streich, 1964). A training system, which the Air Force imaginatively called STP (System Training Program), was embedded in SAGE to train its operators in the mid-1950s. STP included simulated radar inputs, nonradar track inputs, an authoring system for simulation scenarios, a recording capability, and a data reduction feedback feature used for after-action reviews.

STP was a multiplayer training simulation. It served as a prototype and progenitor for a host of military and civilian computer-based simulation systems, as S. R. Mayer (1970), Olsen and Bass (1982), and even Fletcher and Rockway (1986) have suggested. It was the precursor of today's networks of military simulators that engage each other on electronic battlefields (Alluisi, 1991).

Whether STP was a multiplayer game as well as a simulation may depend on the perceptions of its users. Distinctions between games and simulations are varied and a matter for continuing discussion. For the purposes of this chapter, games are considered to be simulations that emphasize engaging, immersing entertainment, often at the expense of realism, in contrast to other simulations, which emphasize realism often at the expense of entertainment. Both games and simulations may involve competition, but that seems more central to games than to simulations.

Multiplayer games have been available for some time, certainly since the advent of multiuser time-sharing systems in the 1960s. Games such as Moonwar[1] and Dogfight hosted on the University of Illinois PLATO system evolved into role-playing games such as Dungeon, which appeared in 1975 and was hosted on time-sharing operating systems supported by PDP-10 mainframes. In about 1978, Dungeon was modified at Essex University in the United Kingdom to support a variety of users playing different roles and was renamed Multi-User Dungeon or MUD (Bartle, 1990). It was immediately and extensively popular.

Personal computing, the Internet, and the World Wide Web have intensified the evolution of games. Even though Dungeon used text commands to control game action, it included line-of-sight displays to simulate what players in different roles might see. Graphics displays in games and simulations continued to evolve with the development of computer display technology. The development of modern graphics-based, role-playing, multiplayer games have been with us since the mid-1990s, when Electronic Arts released Ultima Online. Sony Online's Everquest, released in 1999, may have been the first major hit for multiplayer games, with more than 400,000 subscribers by the end of the year paying $10 per month to play. It was soon followed by Ultima Online, Dark Age of Camelot, Star Wars Galaxies, World of Warcraft, and so on (Robar, 2004). The market for these games continues to grow. By 2005, World of Warcraft had over 1 million players in North America and more than 5 million players worldwide.

Many games and simulations include characters depicted in the action but operating independently under computer control. The quality of the game or simulation depends to a substantial extent on the cognitive behavior and responses of these automated characters. Even when characters operate under some combination of human-computer control, there may be significant cognitive components under computer control, especially when the human operator is managing groups or teams of individual characters. In all cases, the underlying computer capabilities for representing cognition are essential in providing characters that display credible behavior. Many games and simulations include them either implicitly or explicitly.

Games and simulations, multiplayer and otherwise, have significant advantages over other means (e.g., explicit paper-and-pencil testing) for assessing cognitive capabilities (Bennett, Jenkins, Persky, & Weiss, 2003; Drasgow & Mattern, 2005; Fletcher, 2002; Garmine & Pearson, 2006). They employ rich and immersing environments that can take full advantage of the timing, multimedia display, multimodal command, and

[1] Moonwar was lesson 0moon on the PLATO system according to one expert, who claims not have spent any time playing it.

instant scoring capabilities of computer-assisted assessment. They can do so continuously and unobtrusively, and they can yield more detailed and complete representation of what users know and, especially, can do than more traditional assessment methods. They may well take us to a future in which explicit testing is minimized while the extent and depth of assessment is significantly enhanced.

Games, Simulation, and Cognition

Empirical research is investigating the ability of computer games to increase cognitive capabilities (Green & Bavelier, 2003; Hayes, 2005; O'Neil, Wainess, & Baker, 2005; Tobias & Fletcher, in press). Some of these studies indicate that cognitive capabilities can be reliably measured by games themselves. Subrahmanyam and Greenfield (1994) showed that general spatial capabilities such as anticipating targets and extrapolating spatial paths could be assessed in a computer game (Marble Madness). Greenfield, deWinstanley, Kilpatrick, and Kaye (1994) found that general attention skills such as those dealing with requirements for divided visual attention could be assessed (and improved) by playing video games. Hong and Liu (2003) identified three cognitive strategies (trial and error, heuristic, and analogical thinking) used by computer game players and found that these strategies could be used to assess players' expertise.

In discussing cognitive models in simulations and games, Wulfeck (personal e-mail communication, January 17, 2006) asked, "Is there any indication that attention skills or cognitive strategies transfer to any other tasks?" This is a key question. However, research on the transfer of cognitive skills or strategies acquired in games to real-world tasks remains fragmentary, mixed, and scarce.

For instance, Gopher, Weil, and Bareket (1994) found that groups trained for 10 hours on the Space Fortress II game demonstrated performance in a complex and dynamic aircraft flight environment that was superior to an ability-matched control group not exposed to the game. They attributed the success of the game groups to their learning how to cope with attention demands and high cognitive load. On the other hand, Hart and Battiste (1992) found that assigning trainees to an off-the-shelf game (Apache Strike Force), also dealing with flying, had no transfer effects. Fletcher and Tobias (2006) went so far as to suggest that the physical similarities of game and flight conditions do not affect transfer as much as the similarities in attention and cognitive load demands shared by the game and actual flight.

Clearly, the degree to which basic cognitive capabilities acquired and assessed in games and simulations transfer to the performance of real-world

skills and tasks remains to be determined. Fortunately, this chapter avoids this issue by limiting its focus to the representation of cognition in games and simulations in the first place.

The Role of Cognitive Models

Cognitive models concern such processes as human perception, memory, learning, decision making, and, notably, problem solving. They are used to populate games and simulations with synthetic but realistic characters, model the desired end state of learners in instructional applications, and assess the current competencies of users in instructional, decision-aiding, and entertainment applications. Models of current and targeted cognitive end states can help manage users' progress toward achieving instructional objectives, ensure that games and simulations adjust to participant's levels of ability, and provide hints for partial solutions and critiques of completed activities.

In many applications, the performance of participants may be reflected (or overlaid) onto empirically based models of human cognition as the participants interact with games and simulations. This approach has been successfully used in technology-based instruction starting early (e.g., Fletcher, 1975; Sleeman & Brown, 1982) and continuing into the present (Lovett & Anderson, 2005). Use of underlying cognitive models extends ad hoc game and simulation approaches into formal constructs of cognition. Doing so not only yields more comprehensive and generalizable knowledge of players' cognitive processes and abilities but also provides feedback to the models themselves. It will help verify our concepts of human cognition through prediction and exquisitely detailed observation of players' performance. By closing this feedback loop, assessment of cognitive processes and abilities in games and simulations may reveal significant aspects of human cognition that have heretofore been obscured by our limited modalities for assessment. Assessment using games and simulations could transcend its current novelty status by both providing more powerful capabilities for assessing cognitive skills and substantially increasing our understanding of human cognitive processes and their implications for human performance.

Use of games and simulations in assessment is as applicable to teams as to individuals (e.g., Fletcher, 1999; O'Neil, Chung, & Brown, 1997). Members of problem-solving teams must, as Sternberg and Davidson (1992) emphasized, form their own models of the problem that is to be solved. They must also, as Rouse, Cannon-Bowers, and Salas (1992) emphasized, develop a mental model of other participants' knowledge and skill (perhaps by cross training team members in each other's roles and responsibilities) and a shared mental model of the team's goal and subgoal states. In addition, they must develop a shared model of the current situation — a need that has been emphasized by commentators ranging from deGroot (1965) and Chase and Simon (1973) to the current interest in situation awareness (e.g., Endsley, 1995). Finally, teams, and team members, must review the success of their plans, and just as important, they must respond to this feedback by devising new models of the goal, subgoal, and current states. A benefit of using games and simulations to assess team problem solving is that in all these activities mental models and the cognitive processes underlying them are to some extent made visible and explicit through the observable decisions, communications, and actions of team members.

Games and simulations can record all decisions and actions observed from participants' communications, keyboard inputs, and clickstreams. The models of participant knowledge, assumptions, and hypotheses that are inferred from these extensive and detailed records can then be reflected against the "ground truth" known by the system. The outcomes in games and simulations, such as those involving tactics, strategies, and opposing players, may turn out to be unfavorable because of factors over which the participants have no control. If the models are in accord with the true state of the system, then we may be able to assume that participants are taking the right actions for the right reasons, regardless of outcome. By providing a window into the internal cognitive structures and processes a team or team member may be using to solve a problem, cognitive models allow us to distinguish between good problem solving and good luck and between poor problem solving and bad luck.

Modeling Human Behavior

The capabilities of games and simulations to assess cognitive capabilities such as problem solving are embryonic, but the ability to enhance them significantly by using models of human cognition now seems at hand. This work has already begun based on research and development in cognition and efforts to model it, which, in the world of games and simulation, typically comes under the heading of human behavior representation. Modeling is increasingly used in simulations for training and education, analyzing decision alternatives, representing characters and avatars, and designing, developing, and acquiring materiel assets. It appears, implicitly and explicitly, in both military and nonmilitary applications. Its movement into games, multiplayer and otherwise, seems equally at hand.

We are fortunate that a number of systematic reviews and analyses of these models have appeared. Pew and Mavor (1998) reviewed 11 such models, Ritter et al. (2002b) reviewed 7 models not covered by Pew and Mavor, and Morrison (2003) reviewed 19 such models. Morrison's review updated the earlier reports and provided additional analyses.

These models can be implemented in digital form — as computer algorithms. Doing so for any model is a significant demonstration. If the model can be represented in an algorithm, its validity can be tested by comparing its predictions with the observed performance of human participants in games, simulations, and elsewhere. If the model cannot be represented in an effective procedure such as an algorithm, then there is reason to question its adequacy as a model.

Models are especially useful in providing diagnostic as well as summative information. They can provide precise information on individual or team cognitive capabilities for use in devising individually tailored training programs. They can also be used to demonstrate the validity of the model itself, suggesting where it must be modified to account for a more comprehensive range of human cognition. Scientific and technological advances may arise from information of this sort, as well as substantial improvements in our ability to educate, train, and assist learners and users.

Most of these models are systems of if-then, condition-response rules, or *productions*, that simulate cognitive structures and processes. The 19 models Morrison (2003) reviewed are summarized in Table 6.1.

Table 6.2 again summarizes the 19 models by identifying the cognitive functions they cover. Table 6.2 indicates which models, in our judgement, explicitly represent one or more of the following cognitive processes: perception, psychomotor performance, attention, situation awareness, short-term memory, long-term memory, learning, decision making, problem solving, cognitive workload, emotional behavior, and social behavior.

To be indicated in Table 6.2 as present, the documentation and other literature associated with the model had to describe the model's capabilities specifically for that particular function. No inferences were made about the model's potential capabilities to emulate the function with modification. Advocates may reasonably argue that the model could demonstrate additional functions with appropriate modifications. In any case, problem-solving functions are represented in only five of the models. This finding suggests that even though most of these models may be good at reacting to expected situations (i.e., situations for which they are programmed), they may not be so good at adapting to novel, unanticipated situations.

Assessment of Problem Solving

Problem solving and decision making are required when an individual or a group of individuals must achieve a goal, or a combination of competing goals, but are uncertain how to do so (R. E. Mayer & Wittrock, 1996).

Table 6.1 Summary Descriptions of Cognitive Models Reviewed by Morrison (2003)

Model name	Summary description	Reference
Atomic components of thought (ACT) (including atomic components of thought — Rational, ACT-R)	Intended to provide a unified theory of mind and a design basis for instructional environments, intelligent tutors, computer-generated teams and units, and human interfaces. Distinguishes between declarative knowledge (represented with semantic networks) and procedural knowledge (represented using if-then productions). Cognition in ACT-R optimizes choices between environmental demands and computational limitations.	Anderson et al., 2004; Lovett & Anderson, 2005
Adaptive resonance theory (ART)	Family of neural net models designed to explain sensory-cognitive processes (e.g., perception, recognition, attention, reinforcement, recall, and working memory). Postulates bottom-up (e.g., perceptions) and top-down (e.g., expectations, attention control) functions in working memory that interact to produce learning.	Grossberg, 1976a; 1976b Krafft, 2002 http://web.umr.edu/~tauritzd/art
Architecture for procedure execution (APEX)	Intended to reduce time and effort needed to develop models of human performance in complex, dynamic environments such as simulations, explorations of human performance theories, and assessments of equipment design on human performance. Includes goal-directed action selection for tasks and procedures and resource allocation for perceptual (mostly visual), cognitive, and psychomotor functions.	Freed, Dahlman, Dalal, & Harris, 2002 http://www.andrew.cmu.edu/~bj07/apex
Business redesign agent-based holistic modeling system (Brahms)	Models social as well as man-machine interactions. Uses agents to model interactions among physically dispersed groups (e.g., teams) and if-then productions ("detectables" and "beliefs") to model decision making (via "thoughtframes") and behavior within the groups. Emphasizes ethnographic analyses and sociotechnical work practices, activities shaped by sociotechnical environment, and constructivist, situated cognition to model cognition and behavior.	Clancey, Sachs, Sierhuis, & van Hoof, 1998, Acquisti, Clancey, van Hoof, Scott, & Sierhuis, 2001
Cognition and affect project (CogAff) (with associated SimAgent tool kit)	Conceptual space for describing cognitive architectures. Integrates emotional with cognitive processes. Incorporates three layers of cognition (reactive, deliberative, and reflective or meta-cognitive); three layers of information processing (perception, central processing, and action); and three types of emotions (primary based on reaction, secondary based on deliberation, and tertiary based on reflection), all producing different perceptual, memory, and motor functions.	Sloman, 2001, 2003 http://www.cs.bham.ac.uk/~axs/cogaff.html
Cognition as a network of tasks (COGNET) (with associated GINA and iGEN(tool kits)	Intended for cognitive task analysis and description of work domains in multitask environments requiring contemplative, decision-oriented, open-ended responses. Uses three subsystems to represent information processing (sensory/perceptual, mental modeling, action/motor); four forms of if-then, production-based task knowledge (goal-directed task hierarchies, perceptual demons to guide attention, blackboard for organizing declarative information, and possible actions linked to time and resource requirements); and meta-cognitive functions. Allows interfacing with other applications.	Zachary, Campbell, Laughery, Glenn, & Cannon-Bowers, 2001 http://www.chiinc.com/cognethome.shtml

(*continued*)

Table 6.1 Summary Descriptions of Cognitive Models Reviewed by Morrison (2003) (Continued)

Model name	Summary description	Reference
Cognitive complexity theory (CCT) (with associated GLEAN3 tool kit)	CCT is an elaboration of the goal, operators, methods, and selection (GOMS) and the model human processor develop by (Card, Moran, & Newell, 1973). CCT provides a theoretical basis for GOMS and for natural GOMS language. It is focused on human interface design, human-computer interaction, and sequential task performance. Employs device models (transition networks); user models (sequentially executed if-then productions, the fundamental CCT units of cognition, retrieval from long-term memory); and mental operators to represent covert cognitive processes.	Kieras and Polson, 1985; Kieras, 1999
Cognitive objects within a graphical environment (COGENT)	Intended solely to provide tools (via a visual programming environment that evolves with the model built) for cognitive modeling. Assumes functional modularity (cognition as interaction among semiautonomous subsystems) and uses low-level processing components.	Cooper, Yule, & Sutton, 1998; Yule & Cooper, 2000 http://cogent.psyc.bbk.ac.uk
Concurrent activation-based production system (CAPS)	Hybrid model for central cognitive functions (e.g., reading comprehension). Primary focus is on modeling patterns of brain activation patterns in high-level cognition via if-then productions for specific areas of the brain and associative networks for cognitive subsystems. Total activation in working memory is capped, concerned exclusively with declarative knowledge (facts), but with different limits for different individuals. Long-term memory includes procedural and declarative knowledge.	Just & Carpenter, 1992; Just et al., 1999 http://coglab.psy.cmu.edu
Construction-integration theory (C-I theory)	Uses a symbolic theory of sentence comprehension and propositions (actions and objects of the action), stressing goal formation to provide a general model of cognition. Comprehension progresses from approximations to verified integration through mutually reinforced associations and spreading activation in memory. Extended to cover comprehension of novel computer interfaces (LInked model) and new Web sites (CoLiDeS model) and to incorporate concepts from latent semantic analysis (LSA) used to derive meaning from text.	Kintsch, 1998; Landauer & Dumais, 1997; Kitajima & Polson, 1997; Kitajima, Blackmon, & Polson, 2000 http://psych-www.colorado.edu/ics
Distributed cognition (DCOG)	Intended to model individuals' expert behavior with agents that use multiple strategies to respond to a complex environment (air-traffic control). Based on a two-dimensional space: abstraction with three levels (skill-based responses to signals, production-based responses to signs, and knowledge-based responses to symbols) and decomposition (ranging from individual component to total system processing). Processing within this space depends on level of expertise, workload environment, and an individual's preferred level of engagement.	Eggleston, Young, & McCreight, 2000; Eggleston, Young, & McCreight, 2001
Executive process/ interactive control (EPIC)	Intended to model details of peripheral cognitive processes, input (perception) and output (psychomotor responses) to inform human-system interface design by predicting the order and timing of responses. Includes long-term storage of declarative and procedural knowledge and working memory for assessing their application. Capacity and retrieval limitations arise only from perceptual or psychomotor systems, not from central memory store.	Kieras & Meyer, 1995 http://www.eecs.umich.edu/~kieras/epic.html

(*continued*)

Table 6.1 Summary Descriptions of Cognitive Models Reviewed by Morrison (2003) (Continued)

Model name	Summary description	Reference
Human operator simulator (HOS)	Intended to inform human-system interface design by modeling human performance based on the sequence and timing of subtasks organized in networks. Uses simulation objects (configuration of displays and controls), task networks (if-then productions selecting verb-object pairs used to manipulate the objects), and micro-models (times to complete required subtasks involving perception, information processing, and psychomotor responses) to determine human response times.	Wherry, 1976; Harris, Iavecchia, & Dick, 1989; Glenn, Schwartz, & Ross, 1992
Man-machine integrated design and analysis system (MIDAS)	Intended to inform human-system interface design by modeling individuals and interactions among individuals in performing multiple, concurrent tasks. Uses sensory input (operators and perceivable — detectable, recognizable, and identifiable — objects); memory (with declarative — beliefs in long-term memory, contexts in working memory — and procedural components); decision making; attention (with limitations on processing resources); situation awareness (actual and perceived); and psychomotor output to model human operator limitations and capabilities.	Corker & Smith, 1993; Hart, Dahn, Atencio, & Dalal, 2001 http://caffeine.arc.nasa.gov/midas/index.html
Micro systems analysis of integrated network of tasks (Micro Saint)	Simulation tool that uses a detailed task analysis to decompose human performance into a networked hierarchy (with branching logic and sequential dependencies) of discrete tasks and subtasks for which performance estimates can be validated. Network consists of subtask nodes (with launching conditions, time to complete, and effects) and relationships (that may be probabilistic, tactical requiring a threshold value, or multiple initiating more than one subtask). Designed to communicate with other models and applications through middleware.	Laughery & Corker, 1997
Operator model architecture (OMAR) (uses Developers Interface, a graphics tool kit, for developing performance models)	Models human behavior as interactions among independent computational agents representing interacting individuals or cognitive processes within individuals. Allows both sequentially dependent and parallel task performance with order determined by activation levels of tasks — without an explicit executive process. Allows facile interface with other models.	Deutsch, 1998; Deutsch, MacMillan, & Cramer, 1993
PSI	Attempts to integrate motivation with cognitive processes. Based on three levels of needs that interact to determine motive strength and specific goal behaviors: system needs (water and energy); preservation level (pain avoidance); information level (certainty, competence, affiliation). Action strategies first seek automatized skills, then knowledge-based behavior, then trial and error to satisfy goals.	Bartl & Dörner, 1998; Ritter, Baxter, et al., 2002 http://www.uni-bamberg.de/~ba2dp1/psi.html
Situation awareness model for pilot-in-the-loop evaluation (SAMPLE)	Generalized from original effort to model situation awareness of pilots and air crews in air combat. Uses cognitive task analyses, pattern recognition from Klein's recognition-primed decision-making, Endsley's three levels of awareness (detection, identification, and prediction), and Rasmussen's three tiers of action strategy (skill-based pattern recognition, standardized if-then productions, and knowledge-based problem solving) to provide three stages of processing: information processing (with a continuous state estimator and a discrete event detector); situation assessment (with the information fusion and reasoning required by multitasking); and decision making (with a procedure selector and a procedure executor). Output includes information disparity, situation awareness disparity, and combat advantage index.	Rasmussen, 1983; Klein, 1989; Mulgund, Harper, Zacharias, & Menke, 2000

(continued)

Table 6.1 Summary Descriptions of Cognitive Models Reviewed by Morrison (2003) (Continued)

Model name	Summary description	Reference
State, operator, and result (Soar)	Intended as a comprehensive model of human cognition focused on operational task domains depicting all behavior as goal-driven movement through problem spaces that define states and operators for the tasks at hand. Uses a four-cycle iterative process involving input (via human perception); elaboration (matches if-then, condition-action productions in long-term memory with those in working memory to issue proposals for decision making and direct commands for psychomotor actions); output (psychomotor execution); decision (either selects operators or identifies "impasses" requiring a new subgoal until all impasses are resolved). Uses a single process for long-term memory, learning, task representation, and decision making. All learning occurs through "chunking," which occurs through impasse subgoaling and resolution. Emotions arise from situation awareness clarity and confusion. Integrates individual and team knowledge and allows goals and plans to be shared among team members.	Rosenbloom, et al. 1991; Lewis, 2001 http://ai.eecs.umich.edu/soar http://www-2.cs.cmu.edu/afs/cs/project/soar/public/www/home-page.html http://www.isi.edu/soar/soar-homepage.html http://www.nottingham.ac.uk/pub/soar/ nottingham/soar-faq.html http://phoenix.herts.ac.uk/~rmy/cogarch.seminar/soar.html

Table 6.2 Cognitive and Behavioral Functions Represented in Models Reviewed by Morrison (2003)

	Cognitive function represented											
Acronym/abbreviation	Perception	Psychomotor performance	Attention	Situation awareness	Working memory	Long-term memory	Learning	Decision making	Problem solving	Cognitive workload	Emotional behavior	Social behavior
ACT	X	X	X		X	X	X	X	X			
ART	X		X		X	X	X	X				
APEX	X	X				X		X				
Brahms	X	X				X		X				X
CogAff	X	X			X	X		X			X	
COGNET	X	X	X	X	X			X	X	X		
CCT	X	X			X	X		X				
COGENT					X	X	X	X				
CAPS			X		X	X		X	X			
C-I theory			X		X	X		X				
DCOG	X		X		X	X		X		X		
EPIC	X	X			X	X		X				
HOS	X	X	X		X			X				
MIDAS	X	X	X	X	X	X		X				X
Micro Saint	X		X			X		X		X		
OMAR	X		X			X		X				X
PSI	X	X	X		X	X	X	X	X		X	
SAMPLE	X		X	X		X		X				X
Soar	X	X	X	X	X	X	X	X	X		X	X

Note. An X indicates that the function is represented by the model.

They require ingenuity and creativity on the part of the problem solvers to understand the current situation, identify the relevant knowledge and skills they need or possess, and transform them into actions that lead to goal achievement. Decision making focuses on the subsidiary step of identifying alternative paths of action and selecting among them, probably through a recognition-primed process as described by Klein (1989), assessing them through a form of mental simulation (Endsley, 1995; Klein, 1989), and choosing the first workable, or "satisficing" (Simon, 1956), path to pursue. The overall process of evaluating the current situation, generating solution paths, choosing a path, acting on it, and modifying it as needed to meet changing circumstances may be described as *problem solving* (e.g., Miller, Galanter, & Pribram, 1960).

Most real-world problem solving is multivariate, complex, and steeped in uncertainty. It is required in everything from designing menus to deploying military personnel. It is a critical component of the skills needed to ensure workforce readiness and viability in the global marketplace (O'Neil, 1999). Assessing problem solving in environments that resemble the real world as much as possible is a significant undertaking that seems to require some degree of simulation combined with adequate underlying models of both the cognitive processes required in general to solve problems and those used by specific human participants.

Five of the models used to represent human behavior listed by Morrison (2003) and in Table 6.2 specifically address problem solving: atomic components of thought (ACT), cognition as a network of tasks (COGNET), concurrent activation-based production system (CAPS), PSI, and state, operator, and result (Soar). As Morrison emphasized, yes/no judgments do not convey the quality and extent to which these models address any particular function, including problem solving. Accordingly, following is a brief discussion of these five models in more detail. These comments are based on Morrison's review, which provides still more detail.

Atomic Components of Thought

ACT evolved from the human associative memory (HAM) model developed by Anderson and Bower (1973). HAM was a connectionist model of semantic memory that represented Anderson's doctoral research at Stanford University. ACT is a synthesis of HAM and a production system theory of memory (Newell, 1973). The first ACT model appeared in fall 1974, and it has been continually updated since that time. The current model, called ACT-R (R for rational), appeared in the work of Anderson (1993). Updated versions of the ACT-R model have appeared since then (e.g., Anderson et al., 2004).

In distinguishing between knowledge types, ACT refers to declarative and procedural knowledge. It views declarative knowledge as stored information concerning facts about the world. In ACT, this knowledge is modeled as a semantic network, not unlike the memory representation in HAM. In contrast, Anderson contends that our knowledge of actions (i.e., how to do something) is quite different. This procedural knowledge is modeled as a production system. Declarative and procedural knowledge are held in long-term memory. These two systems communicate through

working memory, which is not a separate memory subsystem but rather the subset of knowledge that is currently active.

Problem solving in ACT is accomplished through analogy and example. It may be represented by changes in activation and strength parameters in the semantic network. Productions may not fire because activations levels are below some threshold. Similarly, incorrect productions may fire because their threshold levels are relatively high with respect to correct productions. Activation and strength parameters also affect the latency of responding. A major contribution of ACT is its ability to provide quantitative predictions of performance time and error rates in problem solving.

ACT was originally developed to address cognitive activity and is good at simulating individual intellectual functions, such as problem solving. ACT-R has had several practical applications, including the development of intelligent tutors for math and computer science aimed at secondary education. It has also been used to model human-computer interaction as design aid, and it has provided a framework for interpreting data from brain imaging. However, the ACT-R architecture does not model collective performance, which may be the next step. Anderson has stated that he wants ACT-R to provide computer-generated forces to inhabit training environments and games.

Cognition as a Network of Tasks

COGNET is a symbolic computational model focused on human competency. It was developed by Wayne Zachary in 1989. Its goal is "to facilitate cognitive task analysis and description of specific work domains" (Zachary, Ryder, & Hicinbotham, 1998, p. 16). A more recent version, COGNET-P, was designed to model performance. It includes mechanisms for incorporating time and accuracy constraints and a metacognitive component for control and self-awareness in carrying out cognitive processes, including problem solving (Zachary, Ryder, & Le Mentec, 2002).

COGNET models begin with the assumption that humans are capable of performing multiple tasks simultaneously. COGNET simulates parallel processing with rapid attention switching by maintaining several tasks in various states of completion but allowing only one of these tasks to execute at a time. It is particularly useful for modeling complex time-constrained, multitask situations that require performers to switch the focus of their attention repeatedly as they do in real-world problem solving and decision making.

The COGNET internal information-processing system is intended to apply to all tasks. It is divided into three subsystems: (a) a sensory/perceptual subsystem that converts incoming physical data into symbolic information for use in information processing; (b) an internal cognitive subsystem that constructs and operates a mental model of the world; and (c) an action/motor subsystem for manipulating the external world.

The perceptual and cognitive subsystems are linked by an information store that subsumes short-term, long-term, and working memory. Zachary et al. (2002) do not deny that different short-term and long-term memory effects exist; rather, they contend that the distinction between the two types of memory stores is unnecessary to model cognitive processes.

The COGNET architecture uses a formal production-based system model to represent all forms of task knowledge needed for problem solving. It consists of (a) a network of tasks expressed as goal hierarchies for representing procedural knowledge; (b) perceptual demons that contend ("shriek" with different amplitude) for the attention of the cognitive processor, like Selfridge's (1959) pandemonium model of attention; (c) a "blackboard" for representing and organizing relevant declarative information relating to the problem; and (d) actions for effecting change in the world.

COGNET seems best suited for contemplative, open-ended tasks that are not strongly perceptual-motor in nature. The metacognitive control functions, together with the ability to model independent cognitive agents, should also give COGNET the capability to model coordination among multiple team members. Production-based task knowledge, attention switching, and the metacognitive components all appear to make COGNET well suited for modeling complex cognitive processes such as problem solving.

Concurrent Activation-Based Production System

CAPS was developed by Marcel Just, Patricia Carpenter, and their colleagues at Carnegie Mellon University. The original version of CAPS was a production system for modeling reading, particularly reading comprehension (Thibadeau, Just, & Carpenter, 1982). A unique aspect of the original model was that it incorporated subsymbolic aspects (spreading activation) into the symbolic production system representation.

Development for CAPS is continuing. A more recent version, 4CAPS, is organized into collaborative modules, which are intended to correspond to the functions of different cortical areas (Just, Carpenter, & Varma, 1999). The primary output of 4CAPS is the location and amount of processing per unit time, which is designed to predict the pattern of brain activity recorded by technologies such as functional magnetic resonance imaging and positron emission tomography.

CAPS makes the following assumptions about human cognition:

- Elements in working memory (facts) have value attributes (activations) that reflect their strength or the degree to which they are believed.

- An element can cause a production to fire if it matches the conditions component of the production and the activation value exceeds a specified threshold.
- Cognitive processing is represented by production firings, which cause activation to be propagated. The flow of propagation proceeds from one working memory element (called the source) multiplied by a factor (called the weight) to another element (called the target).
- Processing is explicitly parallel. No limit exists for the number of productions that can fire on the same cycle, and no explicit mechanism exists to resolve conflicts between productions.
- Long-term memory exists separately from working memory.
- The total amount of activation in working memory is capped at some specified value for each individual. This total activation can be used to keep elements active in working memory or to propagate activation by firing productions (Just & Carpenter, 1992).

CAPS models represent central cognitive functions (e.g., comprehension) and do not include peripheral functions, such as perceptual-motor acts. All knowledge is encoded as productions in long-term storage. No mechanisms to acquire or modify that knowledge (i.e., learning) are included in CAPS.

Memory structures and processes are explicitly defined in CAPS. Long-term memory includes procedural and declarative components. Working memory, in contrast, is exclusively declarative: It contains only facts. Forgetting is modeled by decrementing the activation values of "old" elements that remain in working memory from cycle to cycle without receiving explicit activation.

A feature of CAPS is that it represents processing capacity in a theoretically plausible and empirically valid manner. CAPS uses its capabilities to simulate simple problem solving as a comprehension-like process in which the declarative knowledge in working memory is matched with productions from long-term memory. It has been applied to a problem-solving simulation in which a pilot performing a preflight check is interrupted by critical messages that must be comprehended and then either acted on or ignored.

PSI

PSI is usually presented in all capital letters but has not been defined as an acronym. PSI is unique in focusing on the interaction of cognitive, emotional, and motivational processes. It is currently under development by Dietrich Dörner and his colleagues at the Institut für Theoretische Psychologie der Otto-Friedrich-Universität Bamberg. According to Bartl and Dörner (1998), the PSI project is intended to create an intelligent, motivated, emotional agent that can survive in a variety of domains.

The central psychological construct in PSI is motivation. Motivators are portrayed as analogous to tanks filled with liquids, which must be kept within certain tolerance levels. When the level deviates from the ideal, a motivator is launched to activate behaviors to restore the levels.

Bartl and Dörner (1998) sorted six motivators into three categories of needs: (a) system needs (water and energy), intended to sustain an organism's existence; (b) preservation needs (pain avoidance), designed to maintain an organism's structures; and (c) information needs (certainty, competence, and affiliation) with a cognitive or social basis.

Several needs can be active at once. A problem-solving, motive selection mechanism designates a single need as the actual intention. The mechanism selects the intention with the highest expectancy value, which is defined as the product of the perceived probability of fulfilling the need and the level of need. The resulting product is referred to as motive *strength*.

In PSI, emotions and cognitions are not separate processes. There are three primary mechanisms for shaping or modulating cognition and motivation:

- *Activation Level*: The strengths of various needs lead to specific behaviors and to an increase in general activation level, which speeds information processing and may trigger either or both of the two modulators described next.
- *Resolution Level* (RL): Perceptions are modeled as comparisons among schemata. RL refers to the required precision of those comparisons. At low, general activation levels, RL is high, which results in slow but reliable processing. At high activation levels, RL is low, which leads to fast but inaccurate processing.
- *Selection Level* (SL): SL refers to the ability of PSI to change dynamically the threshold needed to activate a need. This mechanism effectively defends intentions against competing needs, thereby protecting PSI against strong behavioral oscillations.

PSI treats memory as a simple log of perceptions and activities and forgetting as a decay of that record. Elements in short-term memory are transitioned in continuous fashion to an episodic memory, and the remnants of the record (stripped of detail) are eventually transitioned into long-term memory. Emotions interact with memory in that records associated with need satisfaction or with pain have a greater chance of passing to long-term memory than do simple sequences of events.

Output from PSI includes momentary states of motives and the speed and accuracy of simulated behavior. According to Ritter, Shadbolt, et al. (2002), PSI can model a wide variety of learning situations, including associative and perceptual learning, operant conditioning, sensory-motor learning, goal learning (i.e., remembering situations that lead to need satisfaction), and aversions (i.e., remembering situations or needs that cause needs). In addition, Ritter, Shadbolt, et al. (2002) reported that PSI includes several built-in problem-solving strategies, including hill climbing and trial and error.

State, Operator, and Result

According to Ritter, Baxter, Avaramides, and Wood (2002), the Soar developer community stopped regarding Soar as an acronym. Hence, it is not usually written in all caps. Soar is perhaps the most popular model gauged by the number of its proponents. As its name implies, Soar's concept of cognition involves a search through a problem space and application of operators to states to achieve a result.

Part of Soar's popularity can be traced to the fact that it is a multifaceted model that addresses disparate audiences. As described by Ritter, Baxter, et al. (2002), Soar provides both a unified theory of cognition and a set of heuristics for developing theories of cognition. It includes principles and constraints from which one can construct applied models of knowledge-based behavior, including problem solving.

Soar has its roots in work begun in the 1950s by Allen Newell, J. C. Shaw, and Herbert Simon to demonstrate that computers could address complex problem solving. The first model produced by this group was the logic theorist (LT), which was designed to devise proofs of geometry theorems (Newell & Simon, 1956; Newell, Shaw, & Simon, 1957). Those same researchers extended the ideas of the LT to different types of problems in their general problem solver model (Newell, Shaw, & Simon, 1958; Newell & Simon, 1972).

Soar is currently under active development at various sites around the world. Some relevant Web addresses are shown in Table 6.1. Soar development is explicitly constrained by three general assumptions about human cognition and behavior: (a) behavior is flexible and goal driven; (b) learning occurs continuously from experience; and (c) elementary cognitive processes occur well within 1 second (Lewis, 2001).

Another guiding principle of Soar is that it should comprise a small set of independent mechanisms (Rosenbloom, Laird, Newell, & McCarl, 1991). This assumption drives the model not only toward simplicity but also toward uniformity in architecture. For instance, Soar uses a single type of process for modeling long-term memory structure, learning, tasks, and decision making.

Soar depicts all behavior as movement through problem spaces. A problem space defines the states and operators that apply to the task at hand. The knowledge required to execute tasks are modeled as productions (condition-action pairs). The conditions define access paths to knowledge stored in memory, whereas the actions define the memory contents themselves (Lewis, 2001).

The course of information processing in Soar is described by the *decision cycle*. Hill (1999) described the decision cycle as a four-phase iterative process:

1. *Input:* Input productions take information from the external world and place the contents into working memory.
2. *Elaboration:* Productions in long-term memory are matched against the contents of working memory and fire in parallel so that all relevant knowledge is retrieved. Productions that fire create proposals for actions that are evaluated in the decision phase and issue direct commands to the motor system.
3. *Decision:* Proposals for action are examined, and as a result, the system selects appropriate operators. If no such action is called for or several competing actions are indicated, then Soar recognizes an impasse, which automatically sets up a subgoal (creates a new space) for resolving the impasse. If the subgoal recognizes another impasse, then another subgoal is declared for solving it, creating a goal stack. This process proceeds in iterative fashion until all impasses have been resolved.
4. *Output:* Motor commands are executed. Resulting changes in internal and external conditions are considered during the decision phase.

In Soar, perception is represented by encoding productions that take data off of a perceptual buffer (called the *input link*) and place the results into working memory. Sensory models are used to filter what information is potentially perceptible.

All long-term memory (procedural, declarative, episodic) is stored as productions. Productions are used not so much to model behavior but to provide content addressable memory. A production's conditions provide associative pathways to contents contained in its action component. Long-term memory is accessed in parallel: All relevant information is retrieved before Soar's decision cycle is completed. Productions can be added to long-term memory through *chunking*, the acquisition of productions that occur during the process of impasse resolution and subgoaling. No procedures exist for deleting productions from long-term memory.

Working memory provides a store of elements that represent the current situation. This mechanism provides the nexus of information processing in Soar. It integrates inputs from the external world, information in productions stored in long-term memory, and results of Soar's internal decision processes. Working memory is not limited in capacity or time.

Soar posits a third type of memory, preference memory, which stores suggestions or imperatives about current operators (Laird, Congdon, & Coulter, 1999).

Preferences are encoded according to fixed semantics, a process that supports the decision stage in Soar information processing and problem solving.

Soar developers extended its capabilities by devising an explicit model of team goals and plans that are shared among team members. The resulting model, called a shell for teamwork (STEAM), represents an integration of team with individual knowledge (Tambe, 1996). STEAM has been used to model coordination among team members in military units and as the underlying method for improving teamwork in RoboCup '97, an international competition to test multiagent systems using soccer as a simulation test bed (Tambe et al., 1999).

An explicit goal of the Soar research program is that the model demonstrate its ability to represent a variety of intelligent behaviors (e.g., Rosenbloom et al., 1991), including toy tasks (e.g., puzzles and games such as Tower of Hanoi and Cryptarithmetic) and more practical domains, such as knowledge-intensive problems in medical diagnosis (Neomycin-Soar), software design (Designer-Soar), and tactical communications (NL-Soar), and learning in complex expert systems (R1-Soar).

Soar has provided creditable models of performance in a variety of complex simulations. Furthermore, its relatively simple architecture fits an impressive range of task domains. Whereas one might argue that ACT-R is more popular among academic users, Soar is more prevalent in operational applications.

Discussion and Conclusions

Simulation is widely and increasingly used to assess the performance and competence of individuals and teams. Games, viewed here as a subset of simulation, are similarly considered as a source of assessment. The intrinsically motivating characteristics of games make them attractive as means to assess mass audiences, as for example a way to assess the preparation of the national voting public to judge policies, legislation, and regulations concerning technology (National Academy of Engineering, in press). Interest in the use of games and simulations in assessment may well increase as techniques develop for the continuous and unobtrusive modeling of abilities from the communication, keyboard, and clickstream interactions of learners and users with technology-based systems (Fletcher, 1999, 2006).

Problems remain, of course. What are the psychometric properties of games and simulations? Is one pass through a game or simulation sufficient for assessment, or are many needed for reliable, valid, precise measurement? How should we identify critical events and decisions? How should we weigh data obtained from the various modes of interactions used in simulations and games? There are more questions of this sort.

Researchers are making progress in this area (e.g., O'Neil, Allred, & Dennis, 1997). One promising approach is based on Mislevy's evidence-centered design (ECD) (Mislevy, Almond, & Lukas, 2003). In ECD, capabilities are identified for some subject area and organized into a graphical framework. ECD then shows how to connect the responses of test-takers working in a complex simulated environment to this framework. Bennett et al. (2003) provided an example of how ECD might be used to assess scientific inquiry skills in a simulation environment. Readers are enthusiastically referred to more complete discussions of these matters by other authors in this book.

Representing human cognitive processes and capabilities should be a key enabler in developing the techniques for assessing problem solving and other cognitive capabilities through the use of games and simulations. The tools, techniques, and frameworks provided by the 19 models discussed by Morrison (2003) and summarized in this chapter provide substantial capabilities, but work remains to be done.

Both the significance and complexity of representing cognitive processes in a credible and practicable manner have been presented to the modeling and simulation community by their own practitioners as a "grand challenge" (Ciancarini et al., 2002).

Giordano, Reynolds, and Brogan (2004) prepared a list of elements required for human behavior representation and identified those that, in their judgment, "cannot be achieved in a tractable manner ... or there is no known way to accomplish them" (p. 915). Their focus was on the capabilities needed to pass Turing's famous Imitation Game, also known as the Turing test (Turing, 1950). Among the items they identified as unachievable are adapting behavior to dynamic scenarios; pattern recognition coupled with appropriate decision making; and complex cognition, reasoning, and learning.

Assessment of problem-solving ability requires an ability to model human cognition — as suggested in this chapter. A fruitful source of these models, again as suggested, is in the efforts to imbue games and simulations with cognitively realistic participants. But, our cognitive modeling goals need not be as ambitious as those targeted by Giordano et al. (2004) for passing the Turing test. To what degree of completeness, then, and which characteristics must cognitive models possess if we wish to obtain adequate assessments of problem-solving ability from games and simulations?

Assessment is usually and properly performed for a reason, usually to inform a decision, for instance, to select individuals for employment, classify individuals into specific job categories, guide progress of individuals and teams toward achieving instructional objectives, certify the readiness of individuals or teams to perform specific tasks, and so forth. The adequacy

of an assessment must depend on its purpose. The same may be said for the adequacy of the underlying model of cognition used in the assessment.

There is evidently a rich assortment of models and modeling capabilities to choose from and adopt in developing assessment of all sorts, including, of course, assessment of problem-solving processes and abilities. Principles for making these choices and ways to adopt them are, we suggest, proper topics for research and development on the use of games and simulations for assessment.

References

Acquisti, A., Clancey, W. J., van Hoof, R., Scott, M., & Sierhuis, M. (2001, December). *Brahms tutorial* (Technical Memorandum TM01-0002, Version 0.9.9.4 RFC). Moffett Field, CA: NASA Ames Research Center.

Alluisi, E. A. (1991). The development of technology for collective training: SIMNET, a case history. *Human Factors, 33*, 343–362.

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*, 1036–1060.

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory*. Washington, DC: Winston.

Bailenson, J. N., Swinth, K. R., Hoyt C. L., Persky, S., Dimov, A., & Blascovich, J. (2005). The independent and interactive effects of embodied agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence: Teleoperators and Virtual Environments, 14*, 379–393.

Bartl, C., & Dörner, D. (1998). *Comparing the behaviour of PSI with human behavior in the BioLab game* (Memorandum Number 32). Bamburg, Germany: Universität Bamburg, Lehrstuhl Psychologie II.

Bartle, R. (1990). *Early MUD history*. Retrieved January 25, 2006 from www.mud.co.uk/richard/mudhist.htm

Bennett, R. E., Jenkins, F., Persky, H., & Weiss, A. (2003). Assessing complex problem solving performances. *Assessment in Education, 10*, 347–359.

Cannon-Bowers, J. A., Salas, E., & Pruitt, J. S. (1996). Establishing the boundaries of a paradigm for decision-making research. *Human Factors, 38*, 193–205.

Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Erlbaum.

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press.

Ciancarini, P., Donohue, G., Lorenz, D., Panzarasa, P., Pritchett, A., Swartout, W., et al. (2002). Simulation and agent oriented software engineering: Working group report of the grand challenges Dagstuhl Seminar. In R. Fujimoto, D. Lunceford, E. Page, & A. M. Uhrmacher (Eds.), *Grand challenges for modeling and simulation–Dagstuhl report* (pp. 62–66). Retrieved May 12, 2005 from http://www.informatik.uni-rostock.de/-lin/GC/report/Dagstuhl_Report.pdf

Clancey, W. J., Sachs, P., Sierhuis, M., & van Hoof, R. (1998). Brahms: Simulating practice for work systems design. *International Journal of Human-Computer Studies, 49*, 831–865.

Cooper, R., Yule, P., & Sutton, D. (1998). COGENT: An environment for the development of cognitive models. In U. Schmid, J. F. Krems, & F. Wysotzki (Eds.), *A cognitive science approach to reasoning, learning, and discovery* (pp. 55–82). Lengerich, Germany: Pabst Science.

Corker, K. M., & Smith, B. R. (1993). An architecture and model for cognitive engineering simulation analysis: Application to advanced aviation automation. In *Proceedings of the AIAA Computing in Aerospace 9 Conference* (pp. 1079–1088). Santa Monica, CA: American Institute of Aeronautics and Astronautics.

deGroot, A. D. (1965). *Thought and choice in chess.* The Hague: Mouton Press.

Deutsch, S. (1998). *Multi-disciplinary foundations for multiple-task human performance modeling in OMAR* (pp. 303–308). Paper presented at the 20th annual meeting of the Cognitive Science Society, Madison, WI.

Deutsch, S. E., MacMillian, J., Cramer, N. L. (1993). *Operator model architecture (OMAR) demonstration final report* (AL/HR-TR-1996-0161). Wright-Patterson Air Forc Base, OH: Armstrong Laboratory, Logistics Research Division.

Drasgow, F., & Mattern, K. (2005). New tests and new items: Opportunities and issues. In D. Bartram & R. Hambelton (Eds.), *Computer-based testing and the Internet: Issues and answers* (pp. 59–76). Chichester, West Sussex, UK: Wiley.

Eggleston, R. G., Young, M. J., & McCreight, K. L. (2000). Distributed cognition: A new type of human performance model. In M. Freed (Ed.), *Simulating human agents: Papers from the 2000 AAAI Fall Symposium* (Technical Report FS-00-03, pp. 8–14). Menlo Park, CA: AAAI Press.

Eggleston, R. G., Young, M. J., & McCreight, K. L. (2001). Modeling human work through distributed cognition. In *Proceedings of the 10th Conference on Computer Generated Forces and Behavioral Representation* (pp. 99–104). Orlando, FL: Simulation Interoperability Standards Organization.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*, 32–64.

Fletcher, J. D. (1975). Models of the learner in computer-assisted instruction. *Journal of Computer-Based Instruction, 3*, 118–126.

Fletcher, J. D. (1999). Using networked simulation to assess problem solving by tactical teams. *Computers in Human Behavior, 15*, 375–402.

Fletcher, J. D. (2002). Is it worth it? Some comments on research and technology in assessment and instruction. In *Technology and Assessment: Thinking Ahead* (pp. 26–39). Washington, DC: National Academy of Science.

Fletcher, J. D. (2006). The ADL vision. In H. F. O'Neill and R. Perez (Eds.), *Web-based learning: Theory, research and practice* (pp. 31–53). Mahwah, NJ: Erlbaum.

Fletcher, J. D., & Rockway, M. R. (1986). Computer-based training in the military. In J. A. Ellis (Ed.), *Military contributions to instructional technology* (pp. 171–222). New York: Praeger.

Freed, M., Dahlman, E., Dalal, M., & Harris, R. (2002). *Apex reference manual for Apex version 2.2.* Moffett Field, CA: NASA ARC.

Garau, M., Slater, M., Bee, S., & Sasse, M. A. (2001). The impact of eye gaze on communication using humanoid avatars. In *Proceedings SIGCHI '01* (pp. 309–316). March 31 to April 4, Seattle, WA.

Garmine, E., & Pearson, G. (Eds). (2006). *Tech tally: Approaches to assessing technological literacy.* Washington, DC: National Academy Press.

Giordano, J. C., Reynolds, P. F., Jr., & Brogan, D. C. (2004). Exploring the constraints of human behavior representation. In R. G. Ingalls, M. D. Rossetti, J. S. Smith, & B. A. Peters (Eds.), *Proceedings of the 2004 Winter Simulation Conference* (pp. 912–920).

Glenn, F., Schwartz, S., & Ross, L. (1992). *Development of a human operator simulator version V (HOS-V): Design and implementation* (Research Note 92-PERI-POX). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Gopher, D., Weil, M., & Bareket, T. (1994). Transfer of skill from a computer game trainer to flight. *Human Factors, 36*, 387–405.

Goldstine, H. H. (1972). *The computer from PASCAL to von Neumann.* Princeton, NJ: Princeton University Press.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*, 534–537.

Greenfield, P. M., deWinstanley, P., Kilpatrick, H., & Kaye, D. (1994). Action video games and informal education: Effects on strategies for dividing visual attention. *Journal of Applied Developmental Psychology, 15*, 195–223.

Grossberg, S. (1976a). Adaptive pattern classification and universal recoding. I: Parallel development and coding of neural feature detectors. *Biological Cybernetics, 23*, 121–134.

Grossberg, S. (1976b). Adaptive pattern classification and universal recoding. II: Feedback, expectation, olfaction, and illusions. *Biological Cybernetics, 23*, 187–202.

Harris, R., Iavecchia, H. P., & Dick, A. O. (1989). The human operator simulator (HOS-IV). In G. R. McMillan, D. Beevis, E. Salis, M. H. Strub, R. Sutton, & L. Van Breda (Eds.), *Application of human performance models to system design* (pp. 275–280). New York: Plenum Press.

Hart, S. G., & Battiste, V. (1992). Flight test of a video game trainer. In *Proceedings of the Human Factors Society 36th Annual Meeting* (pp. 1291–1295). Santa Monica, CA: Human Factors and Ergonomics Society.

Hart, S. G., Dahn, D., Atencio, A., & Dalal, K. M. (2001). *Evaluation and application of MIDAS v2.0* (SAE Technical Paper 2001-01-2648). Warrendale, PA: Society of Automotive Engineers.

Hayes, R. T. (2005). *The effectiveness of instructional games: A literature review and discussion* (Technical Report 2005-004). Orlando, FL: Training Systems Division, Naval Air Warfare Center.

Hill, R. W., Jr. (1999). Modeling perceptual attention in virtual humans. In *Proceedings of the Eighth Conference on Computer Generated Forces and Behavioral Representation* (pp. 563–573). Orlando, FL: Simulation Interoperability Standards Organization.

Hong, J. C., & Liu, M. C. (2003). A study on thinking strategy between experts and novices of computer games. *Computers in Human Behavior, 19*, 245–258.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122–149.

Just, M. A., Carpenter, P. A., & Varma, S. (1999). Computational modeling of high-level cognition and brain function. *Human Brain Mapping, 8*, 128–136.

Kieras, D. E. (1999). *A guide to GOMS model usability evaluation using GOMSL and GLEAN3*. Retrieved August 27, 2002, from the University of Michigan's Electrical Engineering and Computer Science Department site: http://www.eecs.umich.edu/people/kieras/GOMS/GOMSL_Guide.pdf

Kieras, D. E., & Meyer, D. E. (1995). *An overview of the EPIC architecture for cognition and performance with application to human-computer interaction* (EPIC Report No. 5). Ann Arbor: University of Michigan.

Kieras, D. E., & Polson, P. G. (1985). An approach to the formal analysis of user complexity. *International Journal of Man-Machine Studies, 22*, 365–394.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.

Kitajima, M., Blackmon, M. H., & Polson, P. G. (2000). A comprehension-based model of Web navigation and its application to Web usability analysis. In S. McDonald, Y. Waern, & G. Cockton (Eds.), *People and Computers XIV-Usability or Else!* [Proceedings of HCI 2000] (pp. 357–373). Heidelberg, Germany: Springer-Verlag.

Klein, G. A. (1989). Recognition-primed decisions. In W. B. Rouse (Ed.), *Advances in man-machine systems research* (pp. 47–92). Greenwich, CT: JAI Press.

Krafft, M. F. (2002, October). *Adaptive resonance theory*. Retrieved March 5, 2003, from the University of Zurich, Department of Information Technology Web site: http://www.ifi.unizh.ch/staff/krafft/papers/2001/wayfinding/ html/node97.html

Laird, J. E., Congdon, C. B., & Coulter, K. J. (1999). *The Soar user's manual: Version 8.2*. Retrieved June 4, 2000, from http://ai.eecs.umich.edu/soar/docs/manuals/soar8manual.pdf

Laughery, K. R., Jr., & Corker, K. (1997). Computer modeling and simulation. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (2nd ed., pp. 1375–1408). New York: Wiley.

Lewis, R. L. (2001). Cognitive theory, Soar. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (pp. 2178–2183). Amsterdam: Pergamon (Elsevier Science).

Lovett, M. C., & Anderson, J. R. (2005). Thinking as a production system. In K. Holyoak & R. Morrison (Eds.), *Cambridge handbook of thinking and reasoning* (pp. 401–430). New York: Cambridge University Press.

Mayer, R. E., & Wittrock, M. C. (1996). Problem solving transfer. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 47–62). New York: Macmillan.

Mayer, S. R. (1970). Trends in human factors research for military information systems. *Human Factors, 12*, 177–186.

Miller, G. A., Galanter, E. H., & Pribram, K. H. (1960). *The plans and structures of behavior*. New York: Holt, Rinehart, and Winston.

Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2003). *A brief introduction to evidence-centered design* (RR-03-16). Princeton, NJ: Educational Testing Service.

Morrison, J. E. (2003). *A review of computer-based human behavior representations and their relation to military simulations* (IDA Paper P-3845). Alexandria, VA: Institute for Defense Analyses.

Mulgund, S. S., Harper, K. A., Zacharias, G. L., & Menke, T. E. (2000). SAMPLE: Situation awareness model for pilot-in-the-loop evaluation. In *Proceedings of the Ninth Conference on Computer Generated Forces and Behavioral Representation* (pp. 377–378). Orlando, FL: Simulation Interoperability Standards Organization.

Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W. G. Chase (Ed.), *Visual information processing* (pp. 283–310). New York: Academic Press.

Newell, A., Shaw, J. C., & Simon, H. A. (1957). Empirical explorations of the logic theory machine: A case study in heuristics. In *Proceedings of the 1957 Western Joint Computer Conference* (pp. 218–230). Also reprinted in E. A. Feigenbaum and J. Feldman (Eds.), *Computers and thought* (pp. 109–133). New York: McGraw-Hill, 1963.

Newell, A., Shaw, J. C., & Simon, H. A. (1958). Elements of a theory of human problem solving. *Psychological Review, 65*, 151–166.

Newell, A., & Simon, H. A. (1956). The logic theory machine: A complex information processing system. *IRE Transactions in Information Theory, IT-2*, 61–79.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Olsen, J. R., & Bass, V. B. (1982). The application of performance technology in the military: 1960–1980. *NSPI Journal, 41*, 32–36.

O'Neil, H. F., Jr. (1999). Perspectives on computer-based performance assessment of problem solving. *Computers in Human Behavior, 15*, 255–268.

O'Neil, H. F., Jr., Allred, K., & Dennis, R. A. (1997). Validation of a computer simulation for assessment of interpersonal skill. In H. F. O'Neil (Ed.), *Workplace readiness: Competencies and assessment* (pp. 229–254). Mahwah, NJ: Erlbaum.

O'Neil, H. F., Jr., Chung, G. K., & Brown, R. S. (1997). Use of networked simulations as a context to measure team competencies. In H. F. O'Neil (Ed.), *Workplace readiness: Competencies and assessment* (pp. 411–452). Mahwah, NJ: Erlbaum.

O'Neil, H. F., Jr., Wainess, R., & Baker, E. L. (2005). Classification of learning outcomes: Evidence from the computer games literature. *The Curriculum Journal, 16*, 455–474.

Pew, R. W., & Mavor, A. S. (Eds.). (1998). *Modeling human and organizational behavior: Applications to military simulations.* Washington, DC: National Academy Press.

Rasmussen, J. (1983), Skills, rules and knowledge: Signals, signs and symbols and other distinctions in human performance models. *IEEE Transactions on Systems, Man and Cybernetics, SMC-13*, 257–266.

Ritter, F. E., Baxter, G. D., Avaramides, M., & Wood, A. B. (2002). *Soar: Frequently asked questions list.* Retrieved May 29, 2002, from the Pennsylvania State University's Soar Web site: http://ritter.ist.psu.edu/soar-faq/soar-faq.html

Ritter, F. E., Shadbolt, N. R., Elliman, D., Young, R., Gobet, F., & Baxter, G. D. (2002b). *Techniques for modeling human and organizational behaviour in synthetic environments: A supplementary review.* Wright-Patterson Air Force Base, OH: Human Systems Information Analysis Center.

Robar, J. (2004). Multiplayer technology: A primer. *MS&T Magazine, 6*, 26–29.

Rosenbloom, P. S., Laird, J. E., Newell, A., & McCarl, R. (1991). A preliminary analysis of the Soar architecture as a basis for general intelligence. *Artificial Intelligence, 47*, 289–325.

Rouse, W. B., Cannon-Bowers, J. A., & Salas, E. (1992). The role of mental models in team performance in complex systems. *IEEE Transactions on Systems, Man, and Cybernetics, 22*, 1296–1308.

Rowell, J. T., & Streich, E. R. (1964). The SAGE system training program for the Air Force Defense Command. *Human Factors, 6*, 537–548.

Selfridge, O. (1959). Pandemonium: A paradigm for learning. In *Symposium on the Mechanization of Thought Processes* (pp. 511–529). London: HM Stationary Office.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63*, 129–138.

Sleeman, D., & Brown, J. S. (Eds.). (1982). *Intelligent tutoring systems.* New York: Academic Press.

Sloman, A. (2003). How many separately evolved emotional beasties live within us? In R. Trappl, P. Petta, & S. Payr (Eds.), *Emotions in humans and artifacts* (pp. 35–114). Cambridge, MA: MIT Press.

Sternberg, R. J., & Davidson, J. E. (1992). Problem solving. In M. C. Alkin (Ed.), *Encyclopedia of educational research* (Vol. 3, pp. 1037–1045). New York: Macmillan.

Subrahmanyam, K., & Greenfield, P. M. (1994). Effect of video game practice on spatial skills in girls and boys. Special issue: Effects of interactive entertainment technologies on development. *Journal of Applied Developmental Psychology, 15*, 13–32.

Tambe, M. (1996). Teamwork in real-world, dynamic environments. In *Proceedings of the First International Conference on Multi-Agent Systems* (pp. 361–368). Menlo Park, CA: American Association for Artificial Intelligence.

Tambe, M., Adibi, J., Al-Onaizan, Y., Erdem, A., Kaminka, G. A., Marsella, S. C., et al. (1999). Building agent teams using an explicit teamwork model and learning. *Artificial Intelligence, 110*, 215–239.

Thibadeau, R., Just, M. A., & Carpenter, P. A. (1982). A model of the time course and content of reading. *Cognitive Science, 6*, 157–203.

Tobias, S., & Fletcher, J. D. (in press). What research has to say about designing computer games for learning. *Educational Technology.*

Turing, A. M. (1950). Computing machinery and intelligence. *Mind, 236*, 433–460.

Wherry, R. J. (1976). The human operator simulator — HOS. In T. B. Sheridan & G. Johannsen (Eds.), *Monitoring behavior and supervisory control* (pp. 283–293). New York: Plenum Press.

Yule, P., & Cooper, R. (2000, August). *The COGENT tutorial.* Paper presented at the *22nd Annual Conference of the Cognitive Science Society*, Philadelphia, PA.

Zachary, W., Campbell, G. E., Laughery, K. R., Glenn, F., & Cannon-Bowers, J. A. (2001). The application of human modeling technology to the design, evaluation and operation of complex systems. In E. Salas (Ed.), *Advances in human performance and cognitive engineering research* (Vol. 1, pp. 199–247). New York: JAI Press.

Zachary, W. W., Ryder, J. M., & Hicinbothom, J. H. (1998). Cognitive task analysis and modeling of decision-making in complex environments. In J. Cannon-Bowers and E. Salas (Eds.), *Decision-making under stress: Implications for training and simulation* (pp. 315–344). Washington, DC: American Psychological Association.

Zachary, W., Ryder, J., & Le Mentec, J.-C. (2002). *Applied modeling of human competence and performance with COGNET and COGNET-P.* Unpublished manuscript.